MULTIMEDIA UNIVERSITY

STUDENT ID NO

MULTIMEDIA UNIVERSITY

FINAL EXAMINATION

TRIMESTER 2, 2017/2018

BSA1024 – STATISTICS

(All sections / Groups)

12 MARCH 2018
9.00 a.m. - 11.00 a.m.
(2 Hours)

INSTRUCTIONS TO STUDENTS

1. This question paper consists of NINE (9) printed pages inclusive of the cover page, formulae sheet and statistical tables.

2. This question paper consists of **FOUR** structured questions. Attempt **ALL** questions.

3. Students are allowed to use non-programmable scientific calculators with no restrictions. Statistical tables are attached at the end of the question paper.

4. Please use **pen** to write the answers.

5. Please print all your answers in the **Answer Booklet** provided.

Question 1 [Total = 25 Marks]

a) The "expense ratio" is a measure of the cost of managing the portfolio. Investors prefer a low expense ratio, all else equal. Below are the expense ratios for 14 randomly selected stock funds.

1.12	1.44	1.27	1.75	0.99	1.45	1.19
0.60	2.10	0.73	0.90	1.79	1.35	1.08

(i) Determine the mean and standard deviation for the expense ratio. [2 marks]

(ii) Determine whether there is any outlier in the sample. [13 marks]

(iii) Determine the coefficient of variation for the expense ratio for the stock funds. [2 marks]

b) A computer centre has three printers, A, B and C, which print at different speeds. Programs are routed to the first available printer. The probability that a program is routed to printer A, B and C are 0.6, 0.3 and 0.1 respectively. Occasionally a printer will jam and destroy a printout. The probability that printers A, B and C will jam are 0.01, 0.05 and 0.04 respectively. Your program is destroyed when printer jams.

(i) Find the probability that your program is destroyed. [5 marks]

(ii) Find the probability that those jams to be caused by printer B. [3 marks]

Question 2 [Total = 25 Marks]

a) Experience indicates that 8% of men's trousers dropped off for dry cleaning will have an object in the pocket that should be removed before cleaning. Suppose that 10 pairs of trousers are dropped off at the dry cleaning shop,

(i) find the probability that none of the trousers has an object in the pocket? [2 marks]

(ii) find the probability that at most 3 pairs of trousers has an object in the pocket? [7 marks]

(iii) find the mean and standard deviation for the number of trousers without object in the pocket. [4 marks]

b) The length of a time-out during televised professional football game is normally distributed with a mean of 1.4 minutes and a standard deviation of 10 seconds.

(i) If the network runs consecutive commercials totalling 80 seconds, what is the probability that the play will resume before the commercials are over? [3 marks]

(ii) What is the probability of 85 to 105 seconds of time-out? [5 marks]

(iii) Find the maximum length of time-out for 90% of the game? [4 marks]

Continued...

Question 3 [Total = 25 Marks]

a) A ski outfit company in Vail owns two shops, one on the east side and one on the west side. Sales data showed that at the eastern location, there were 56 ski outfits in pink sold out of 304 ski outfits sold. At the western location, there were 145 ski outfits in pink sold out of 562 ski outfits sold.

(i) Can we conclude that ski outfits in pink have more demand in the western location than the eastern location? [14 marks]
(ii) Find the p-value for the test. [3 marks]

b) The hotel industry is very interested in understanding how tourists spend money. In order to measure the price and demand changes in few important components of a tourist's budget, a statistician computed the average cost and demand of a hotel room (one night), a meal and rented car in 2000 and in 2008. The results of these computations are shown in the accompanying table.

Item	**Year 2000**		**Year 2008**	
	Price	**Quantity**	**Price**	**Quantity**
Hotel (night)	85	14	120	24
Meal	13	18	20	15
Rented car	25	10	45	30

(i) Compute and interpret Laspeyres price index for 2008 using 2000 as the base period. [4 marks]
(ii) Compute and interpret Paasche price index for 2008 using 2000 as the base period. [4 marks]

Question 4 [Total = 25 Marks]

The general manager of a chain of furniture stores believes that experience is the most important factor in determining the level of success of a salesperson. To examine this belief she records last month's sales (in $1,000s) and the years of experience of 10 randomly selected salespeople. Partial of the output are presented below.

Continued...

ANOVA			
	df	*SS*	*MS*
Regression	1	406.7203	406.7203
Residual	8		2.4725
Total	9		
	Coefficients	*Standard Error*	
Intercept	8.63	0.8517	
Experience	1.0817	0.0843	

a) Determine the least squares regression line. [3 marks]
b) Interpret the slope coefficient. [4 marks]
c) Find the coefficient of determination and interpret the value. [6 marks]
d) If a sales person has 15 years of experience, how much he is expected to sell? [3 marks]
e) Test whether experience has significant relationship with the sales. [9 marks]

End of Questions

STATISTICAL FORMULAE

A. DESCRIPTIVE STATISTICS

Mean: $\overline{X} = \dfrac{\sum_{i=1}^{n} X_i}{n}$

Standard deviation: $s = \sqrt{\dfrac{\sum_{i=1}^{n} X^2}{n-1} - \dfrac{\left(\sum_{i=1}^{n} X_i\right)^2}{n(n-1)}}$

Coefficient of Variation: $CV = \dfrac{s}{\overline{X}} \times 100\%$

Pearson's Coefficient of Skewness: $S_k = \dfrac{3(\overline{X} - Median)}{s}$

B. PROBABILITY

Basic Probability Rules

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$P(A \cap B) = P(A) \cdot P(B)$ if A and B are independent

$P(A \mid B) = P(A \cap B) \div P(B)$

Binomial Probability Distribution

If X follows a Binomial distribution, B(n, p) where $P(X = x) = {}^{n}C_{x} p^{x} q^{n-x}$

then the mean = $E(X) = np$ and variance = $VaR(X) = npq$ where $q = 1 - p$

Poisson Probability Distribution

If X follows a Poisson distribution, Po(λ) where $P(X = x) = \dfrac{e^{-\lambda} \lambda^{x}}{x!}$

then the mean = $E(X) = \lambda$ and variance = $VaR(X) = \lambda$

Normal Probability Distribution

If X follows a Normal distribution, N(μ, σ) where the mean = $E(X) = \mu$ and variance = $VaR(X) = \sigma^2$ then $Z = \dfrac{X - \mu}{\sigma}$

C. CONFIDENCE INTERVAL ESTIMATION AND SAMPLE SIZE DETERMINATION

$(100-\alpha)\%$ confidence interval for population mean (σ known): $\overline{X} \pm Z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$

$(100-\alpha)\%$ confidence interval for population mean (σ unknown): $\overline{X} \pm t_{\alpha/2,n-1}\left(\frac{s}{\sqrt{n}}\right)$

$(100-\alpha)\%$ confidence interval for population proportion: $p \pm Z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}$

Sample size determination for population mean: $n \geq \frac{(Z_{\alpha/2})^2\sigma^2}{E^2}$

Sample size determination for population proportion: $n \geq \frac{(Z_{\alpha/2})^2 p(1-p)}{E^2}$

Where E = limit of error in estimation

D. HYPOTHESIS TESTING

One sample mean test:

Standard deviation σ known	Standard deviation σ unknown
$Z = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$	$t = \frac{\overline{X}-\mu}{s/\sqrt{n}}$

One sample proportion test:

$$Z = \frac{p-\pi}{\sqrt{\frac{\pi(1-\pi)}{n}}}$$

Two sample mean test:

Standard deviation σ known	Standard deviation σ unknown
$Z = \frac{\overline{X}_1-\overline{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$	$t = \frac{\overline{X}_1-\overline{X}_2}{\sqrt{S_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$ where $S_p^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{(n_1-1)+(n_2-1)}$

Two sample proportion test:

$$Z = \frac{p_1-p_2}{\sqrt{\overline{p}(1-\overline{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \text{ where } \overline{p} = \frac{n_1p_1+n_2p_2}{n_1+n_2} = \frac{X_1+X_2}{n_1+n_2}$$

E. REGRESSION ANALYSIS

Correlation coefficient:

$$r = \frac{\sum XY - \left[\frac{\sum X \sum Y}{n}\right]}{\sqrt{\left[\sum X^2 - \frac{\left(\sum X\right)^2}{n}\right]\left[\sum Y^2 - \frac{\left(\sum Y\right)^2}{n}\right]}} = \frac{COV(X,Y)}{\sigma_X \sigma_Y}$$

Simple linear regression:

Population model: $Y = \beta_0 + \beta_1 X_1 + \varepsilon$

Sample model: $\hat{y} = b_0 + b_1 X_1 + e$

ANOVA table for linear regression:

Source	Degrees of freedom	Sum of squares	Mean squares
Regression	p	SSR	MSR = SSR/p
Error	$n-p-1$	SSE	MSE = SSE/$(n-p-1)$
Total	$n-1$	SST	

Test statistic for significance of the overall regression model: $F = MSR / MSE$

Test statistics for significance of the predictor variable

$t_i = \frac{b_1 - \beta_1}{S_{b_i}}$ and the critical value $= \pm t_{\alpha/2, n-p-1}$

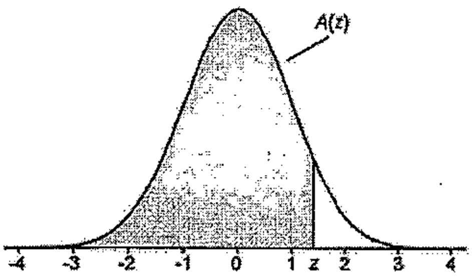
F. INDEX NUMBERS

Simple price index $P = \frac{p_t}{p_0} \times 100$	Aggregate price index $P = \frac{\sum p_t}{\sum p_0} \times 100$
Laspeyres price index (LPI) $P = \frac{\sum p_t q_0}{\sum p_0 q_0} \times 100$	Paasche price index (PPI) $P = \frac{\sum p_t q_t}{\sum p_0 q_t} \times 100$
Fisher's ideal price index $F = \sqrt{LPI \times PPI}$	Value index $V = \frac{\sum p_t q_t}{\sum p_0 q_0} \times 100$

STATISTICAL TABLES **1**

TABLE A.1

Cumulative Standardized Normal Distribution

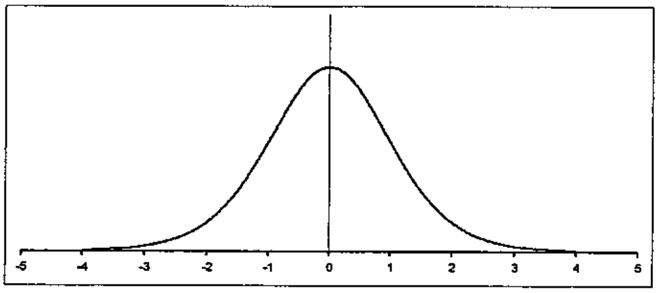


$A(z)$ is the integral of the standardized normal distribution from $-\infty$ to z (in other words, the area under the curve to the left of z). It gives the probability of a normal random variable not being more than z standard deviations above its mean. Values of z of particular importance:

z	$A(z)$	
1.645	0.9500	Lower limit of right 5% tail
1.960	0.9750	Lower limit of right 2.5% tail
2.326	0.9900	Lower limit of right 1% tail
2.576	0.9950	Lower limit of right 0.5% tail
3.090	0.9990	Lower limit of right 0.1% tail
3.291	0.9995	Lower limit of right 0.05% tail

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999							

Table 2: The t-distribution ($t_{\alpha, df}$)



$\alpha =$	0.10	0.05	[redacted]	0.01	0.005	0.001	0.0005
df = 1	3.078	6.314	12.706	31.821	63.657	318.309	636.619
2	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
[redacted]	1.397	1.860	[redacted]	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	1.296	1.671	2.000	2.390	2.660	3.232	3.460
120	1.289	1.658	1.980	2.358	2.617	3.160	3.373
∞	1.282	1.645	1.960	2.326	2.576	3.090	3.291